Single-molecule localization microscopy:

Data analysis

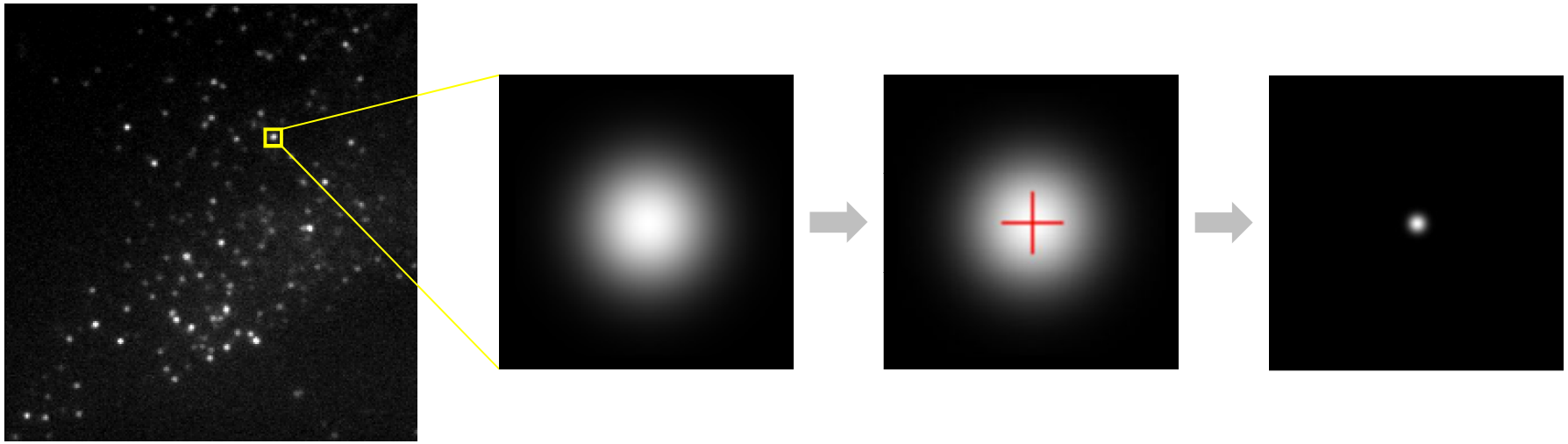
EMBL Advanced Course:
Super-Resolution Microscopy

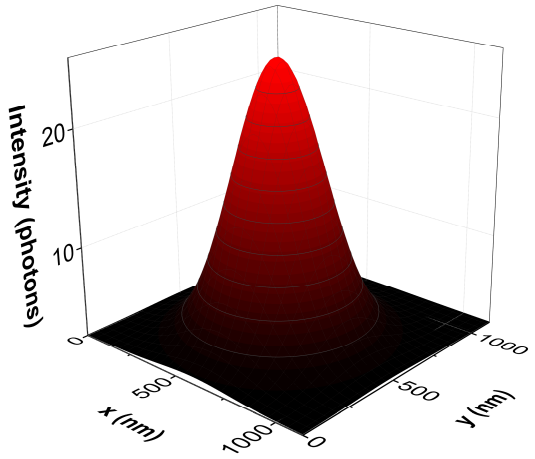
20 - 25 Jul 2015

Franziska Fricke & Mike Heilemann

Single-Molecule Biophysics
Goethe-University of Frankfurt

Single-molecule localization microscopy

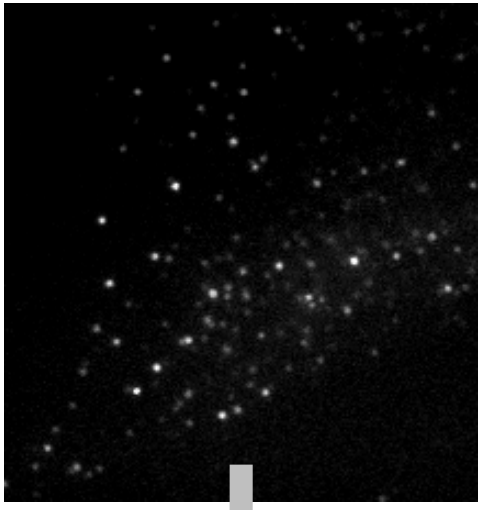




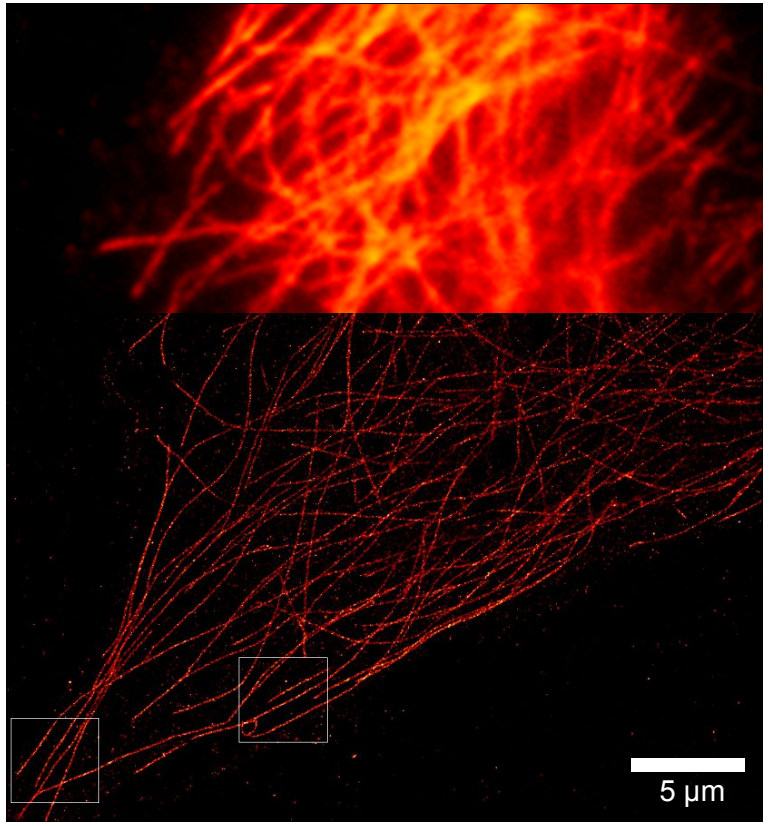
Fit Model: 2D Gaussian function

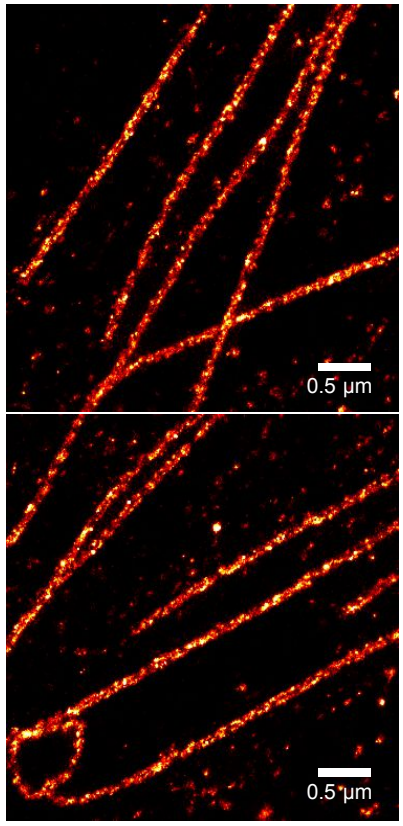
$$I(x,y) = I_0 + \frac{N}{2\pi\sigma_x\sigma_y} exp\left(-\left(\frac{x-x_c}{\sqrt{2}\sigma_x}\right)^2 - \left(\frac{y-y_c}{\sqrt{2}\sigma_y}\right)^2\right)$$

Reconstructing the image



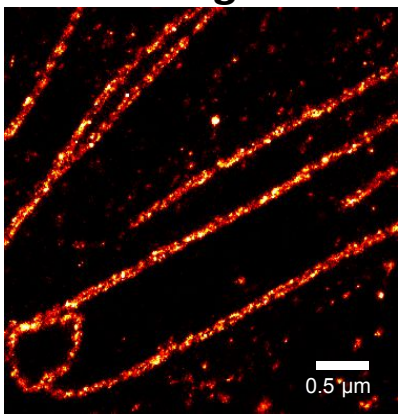
x [nm]	y [nm]	t [ms]	I [Photons]
12.148	8.876	0	2846
68.475	97.829	20	1861
23.073	58.721	20	1770
84.058	63.144	40	1023
200.702	28.336	40	1136
315.832	105.179	40	944
35.063	11.776	60	978
...	...	...	...





Software

image



RapidStorm

Analyzing blinking movies (.tif)

Output: Image and localization table

localization table

x [nm]	y [nm]	t [ms]	I [Photons]
12.148	8.876	0	2846
68.475	97.829	20	1861
23.073	58.721	20	1770
84.058	63.144	40	1023
200.702	28.336	40	1136
315.832	105.179	40	944
35.063	11.776	60	978
…	…	…	…

Fiji

*Quantitative analysis of **image***

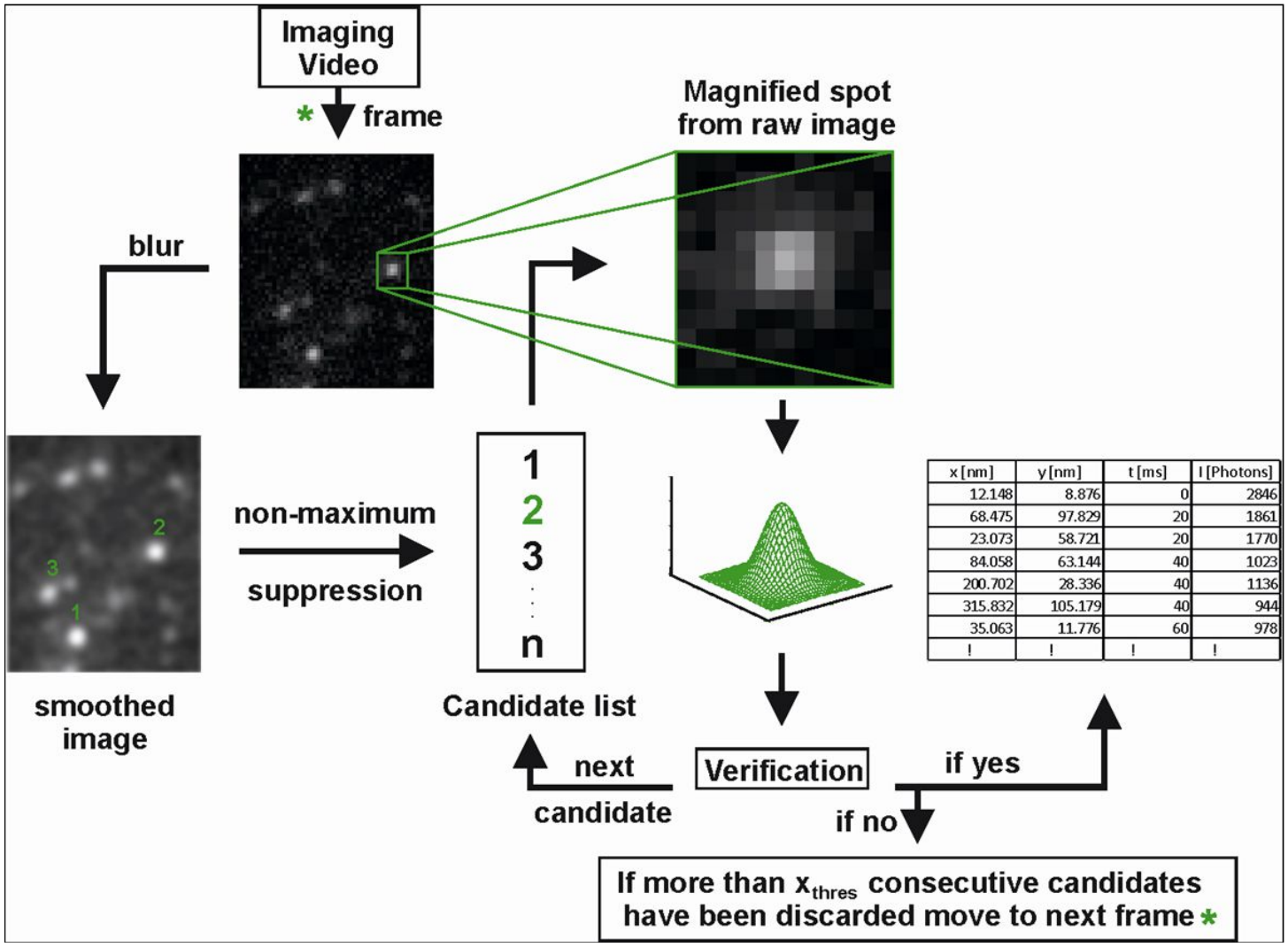
- Resolution estimation
- Image-based cluster analysis
- …

Lama

*Quantitative analysis of **localization table***

- Localization precision
- Coordinate-based cluster analysis
- …

RapidStorm localization routine



Gaussian function

1D Gaussian:

$$I(x) = \frac{N}{\sqrt{2\pi}\sigma} e^{-\left(\frac{x-x_c}{\sqrt{2}\sigma}\right)^2}$$

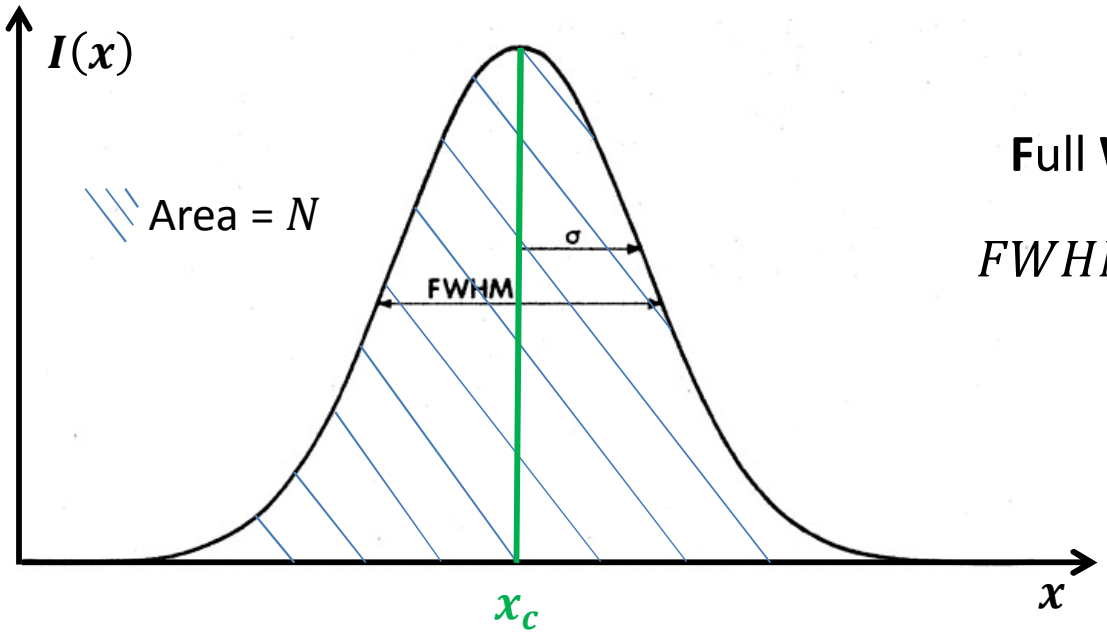
Normal distribution = normalized Gaussian

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\left(\frac{x-x_c}{\sqrt{2}\sigma}\right)^2} dx = 1 \quad \rightarrow \quad \int_{-\infty}^{\infty} I(x)dx = N$$

N = number of photons

σ = standard deviation

x_c = fluorophore position



Full **W**idth at **H**alf **M**aximum:

$$FWHM = 2\sqrt{2ln2}\sigma \cong 2.35\sigma$$

Localization Precision

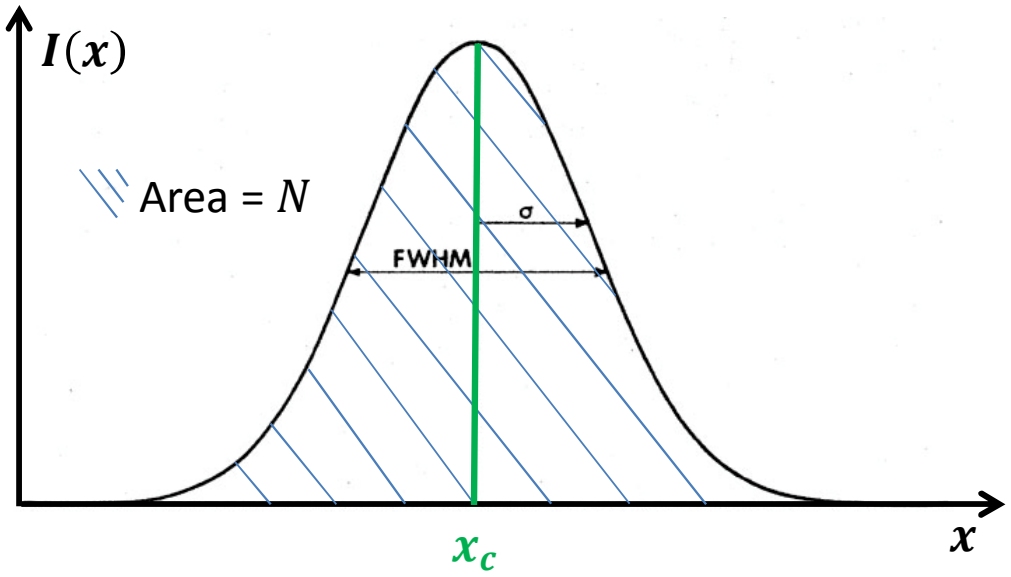
How good can we determine the position of a fluorophore?

Error Δx in position x_c

- 1 photon: $\Delta x = SEM = \sigma$
- N photons: $\Delta x = SEM = \frac{\sigma}{\sqrt{N}}$

SEM = Standard Error of the Mean

(N photons equivalent to N position measurements)



R. Thompson et al. *Biophys J* **82**, 2775-83 (2002)

Localization Precision

How good can we determine the position of a fluorophore?

Further errors are introduced by pixelation noise and background:

Mortensen:

$$\Delta x = \sqrt{\frac{\sigma^2 + a^2/12}{N}\left(\frac{16}{9} + \frac{8\pi(\sigma^2 + a^2/12)b^2}{a^2 N}\right)}$$

a = pixel size

b = background noise

Typical values for error Δx:

Alexa Fluor 647

$\sigma = 150$ nm $\quad (FWHM \cong \lambda/2 \cong 350$ nm)

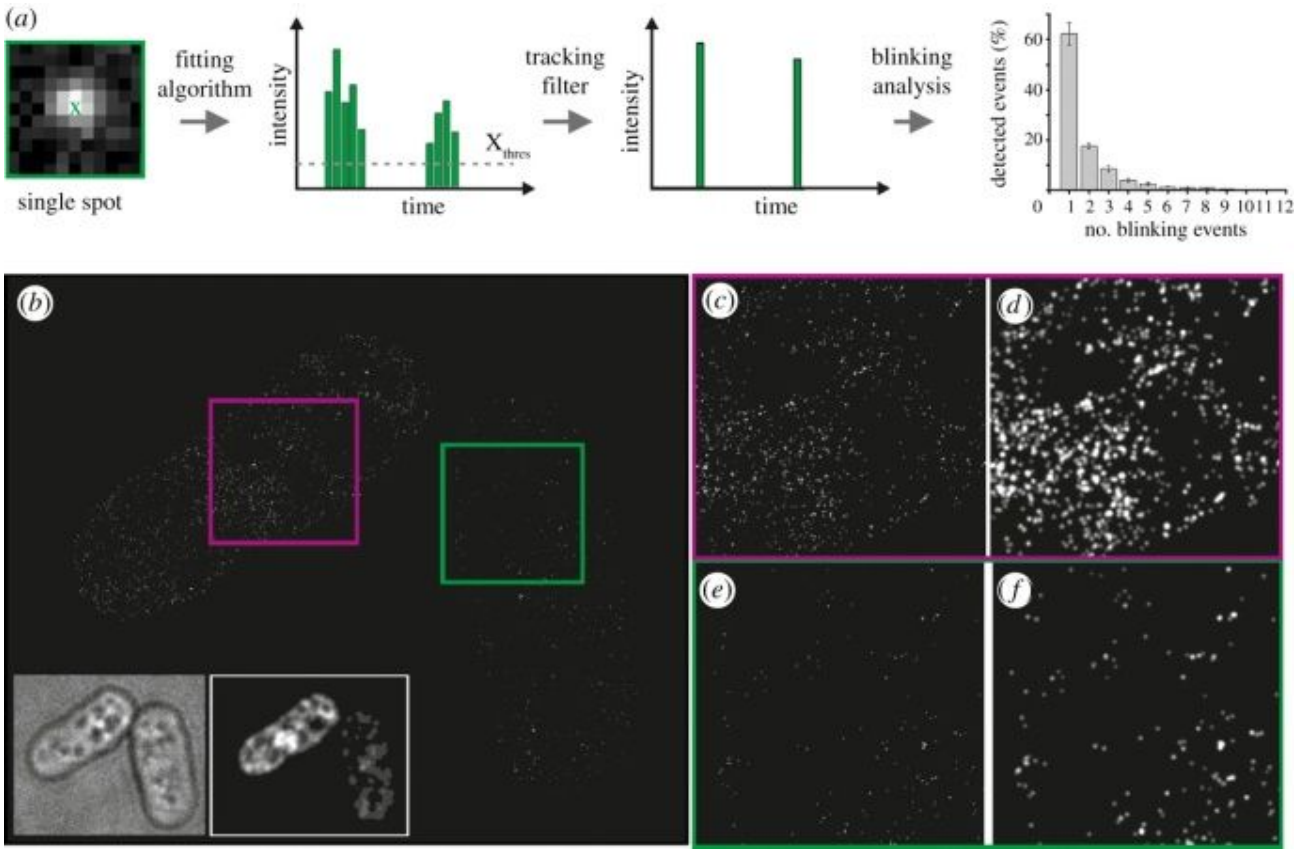
$N = 1000$ photons; $\quad a = 160$ nm; $\quad b = 5$ photons

$$\Delta x_{Mortensen} = 7.7\ nm$$

K. Mortensen et al. *Nat Meth* **7**, 377-81 (2010)

Blinking of fluorophores

Grouping of emissions that last for multiple frames



Important parameters:

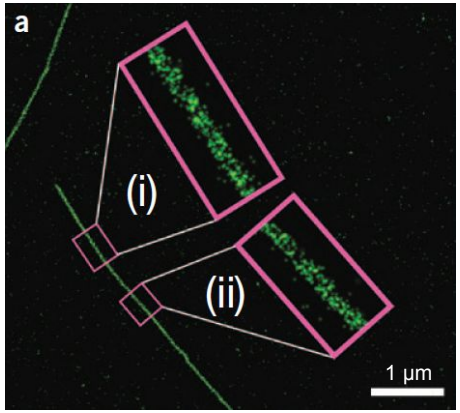
- grouping radius (e.g. multiples of localization precision Δx)
- Grouping time: number of allowed „dark frames“ for fluorophore

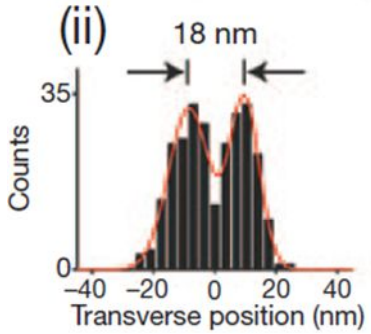
D. Lando et al. *Open Biol* (2012)

Experimental Localization Precision

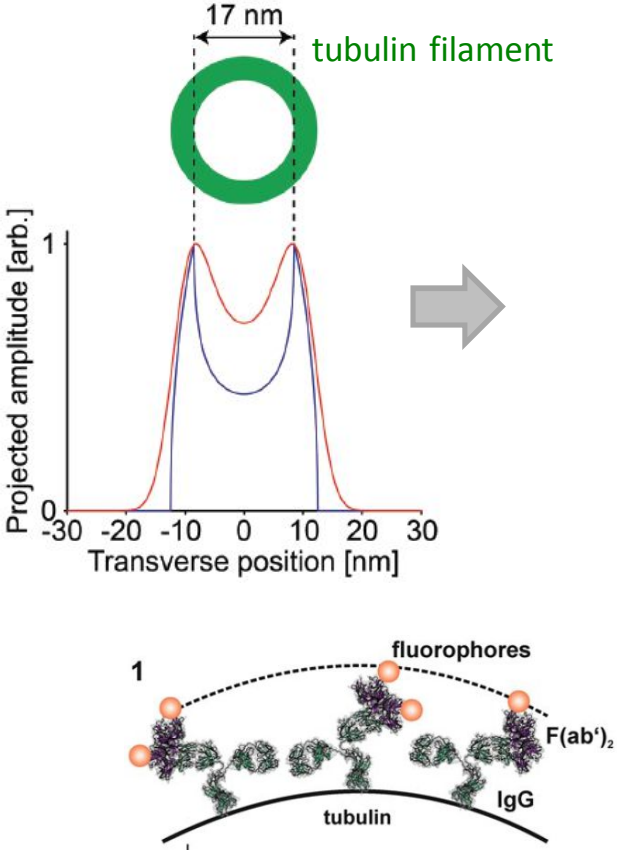
Image resolution = FWHM of a microtubule?

microtubules polymerized and labeled *in vitro*

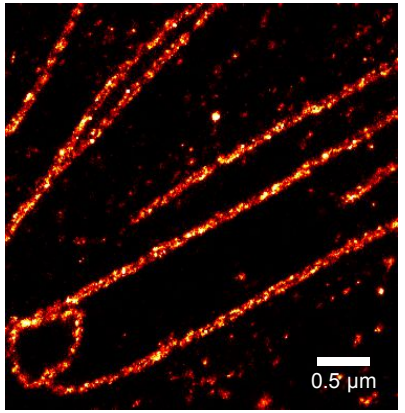


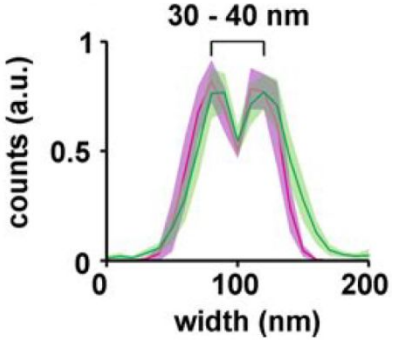


Localization precision ≈ 2 nm



Immunostained microtubules of U2OS



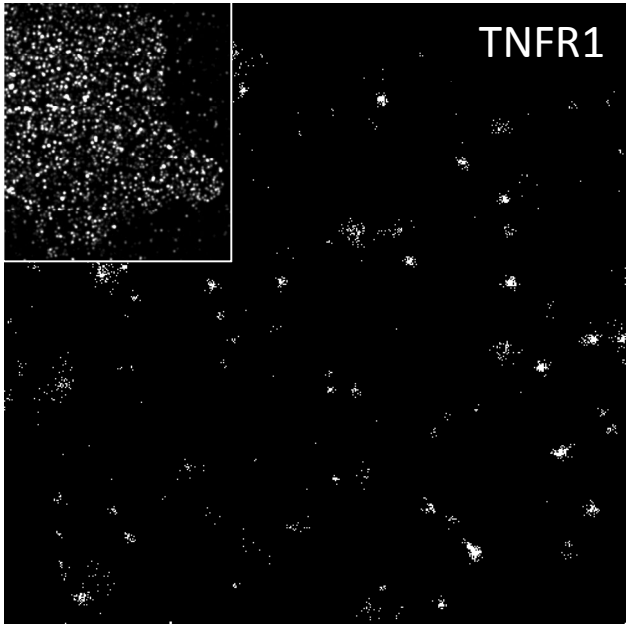


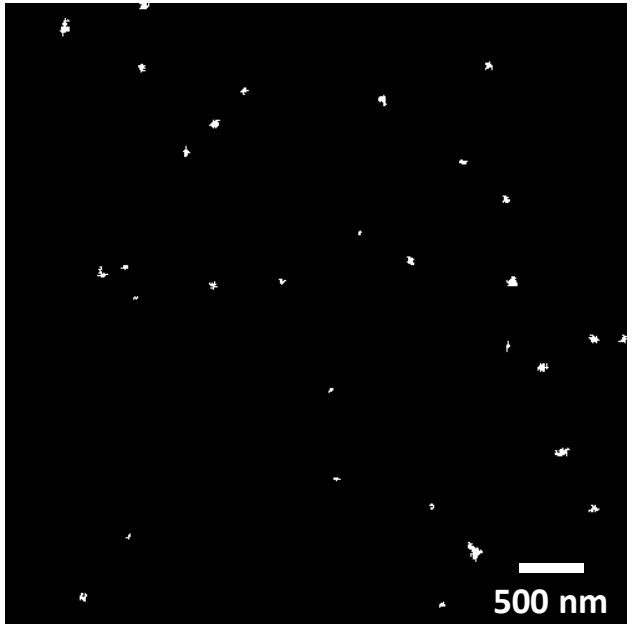
Localization precision ≈ 9 nm

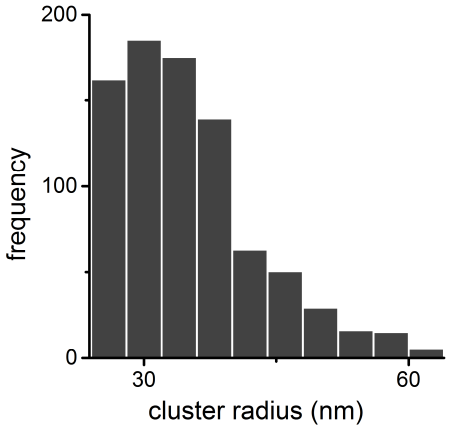
U. Endesfelder et al. *Nat Meth* **11**, 235-8 (2014);
U. Endesfelder et al. *Histochem Cell Biol* **141**, 629-38 (2014); J. Vaughan et al. *Nat Meth* **9**, 1181-4 (2012)

Morphological cluster analysis

Analyzing protein accumulations







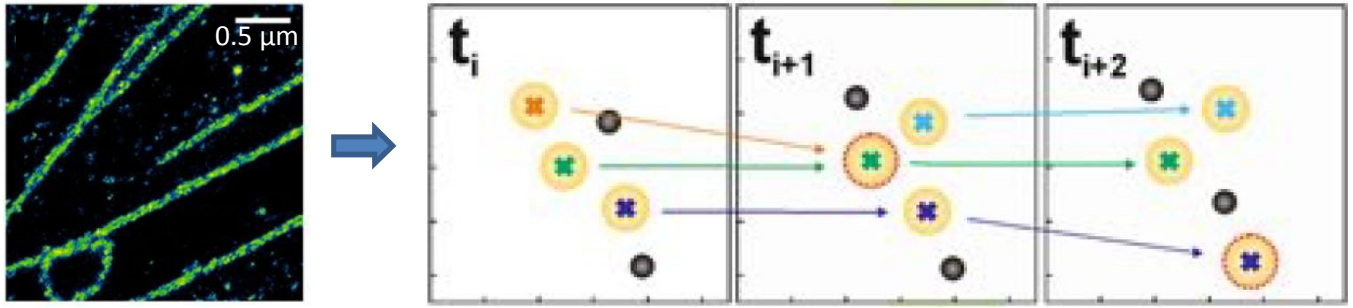
- Density: clusters per area
- cluster size
- localizations per cluster
- …

Experimental Localization Precision

How good can we determine the position of a fluorophore?

Analysis based on nearest neighbor localizations in adjacent frames

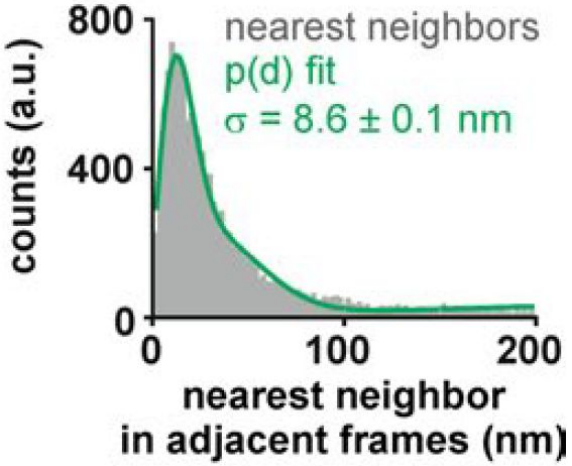
- Determine nearest neighbor distance distribution in adjacent frames



- Fit distribution p(r) to obtain $\Delta x = \sigma$

Experimental:

$$p(r) = \frac{r}{2\sigma^2} e^{-\left(\frac{r}{2\sigma}\right)^2} + \text{corr(Gauss)} + \text{corr(linear)}$$

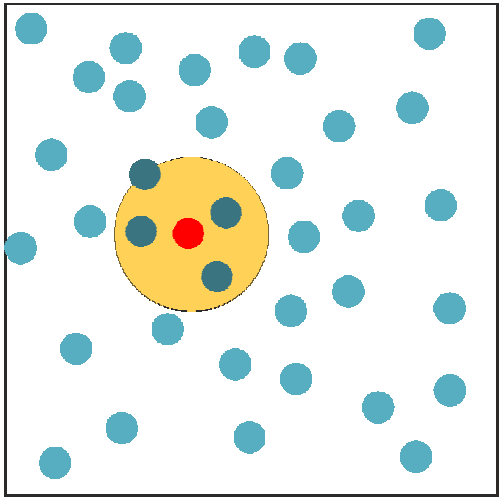


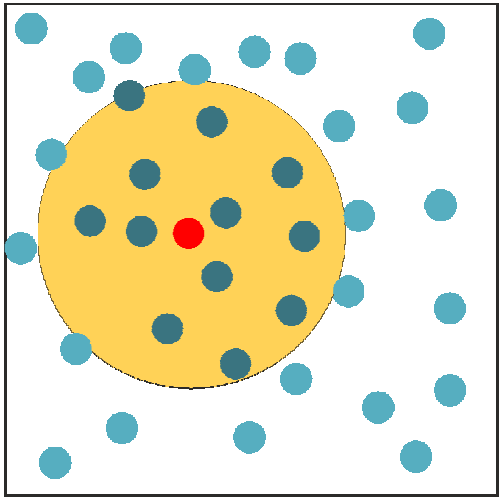
U. Endesfelder et al. *Histochem Cell Biol* **141**, 629-38 (2014)

Ripley's functions

How can we distinguish spatial inhomogeneity?

Ripley's K, L and H functions: *Compare point pattern with uniform point distribution*





N = number of localizations in ROI

A = size of ROI

p_i = localization i (here: red point)

N_{p_i} = number of locs. around p_i within distance $d \leq r$

Ripley's K function:

$$K(r) = \frac{A}{N^2} \sum_{i=1}^{N} N_{p_i} \, (d \leq r)$$

B. D. Ripley, *J Appl Probability* **13**, 255-66 (1976)

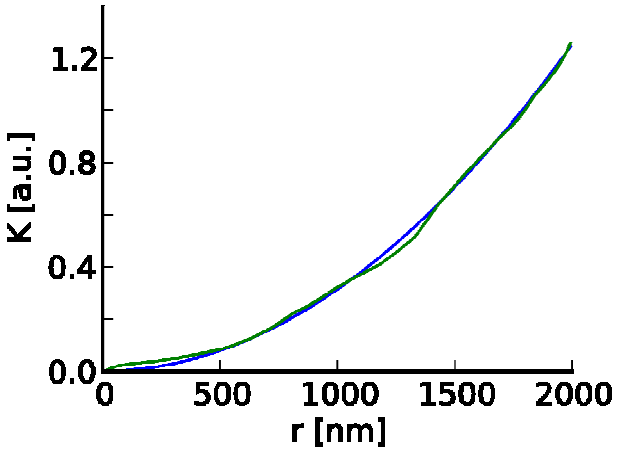
Ripley's functions

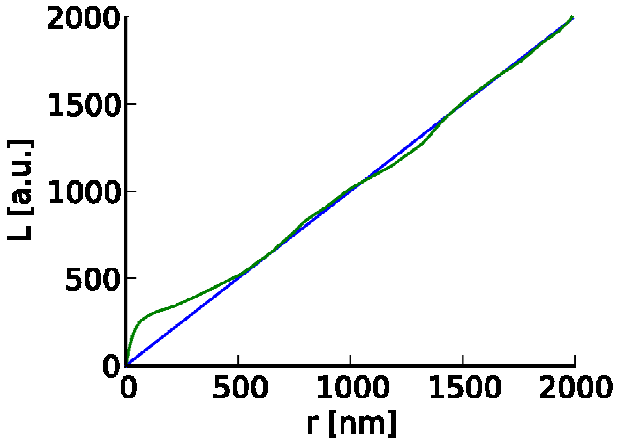
How can we distinguish spatial inhomogeneity?

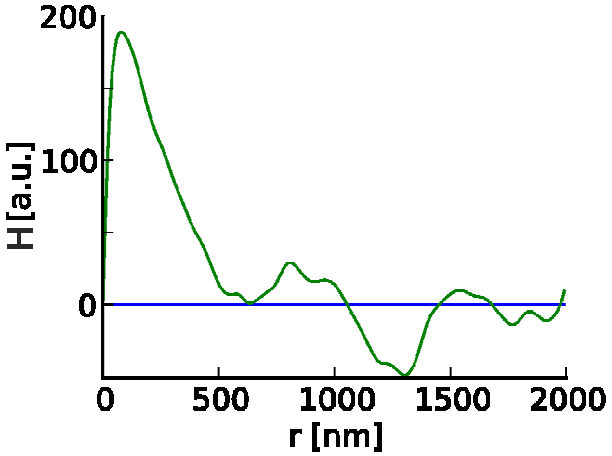
$$K(r) = \frac{A}{N^2} \sum_{i=1}^{N} N_{p_i} \, (d \leq r)$$

$$L(r) = \sqrt{\frac{K(r)}{\pi}}$$

$$H(r) = L(r) - r$$







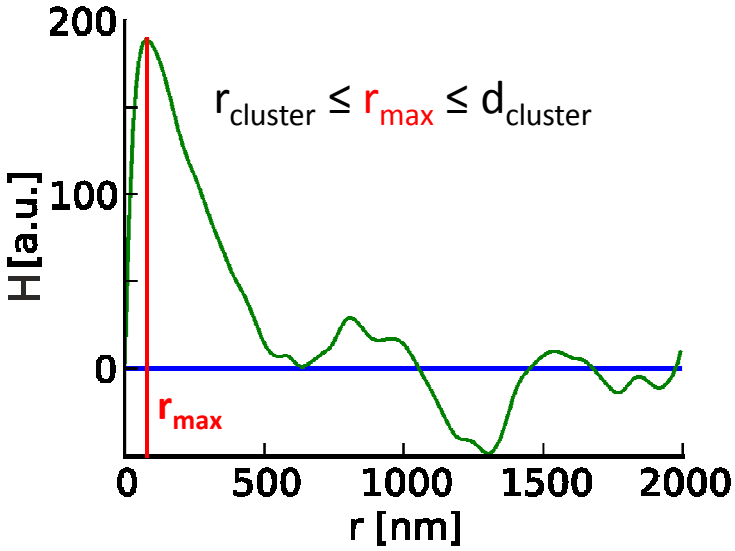
- uniform: $\boldsymbol{K(r) = \pi r^2}$
- clustered pattern

- uniform: $\boldsymbol{L(r) = r}$
- clustered pattern

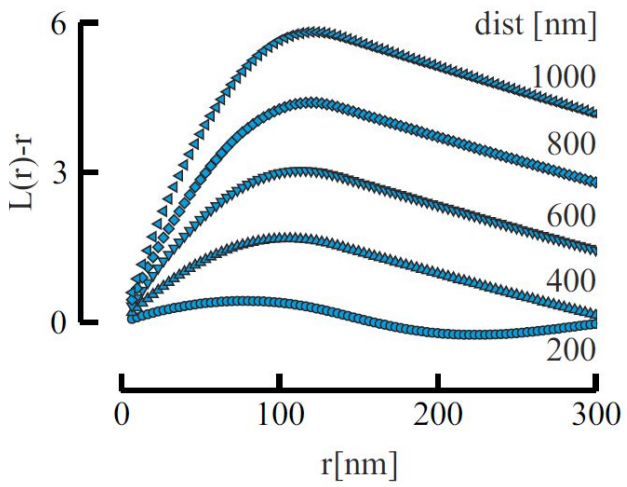
- uniform: $\boldsymbol{H(r) = 0}$
- clustered pattern

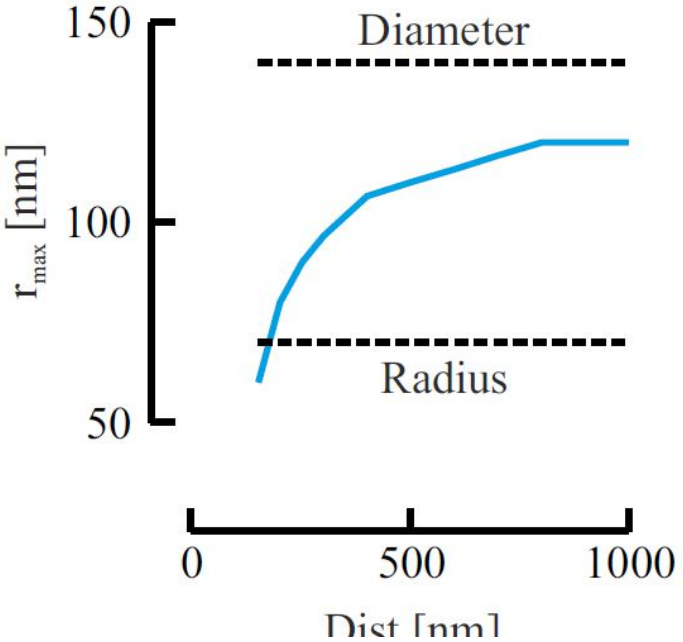
Ripley's functions

Domain size



- uniform: $\boldsymbol{H(r) = 0}$
- clustered pattern

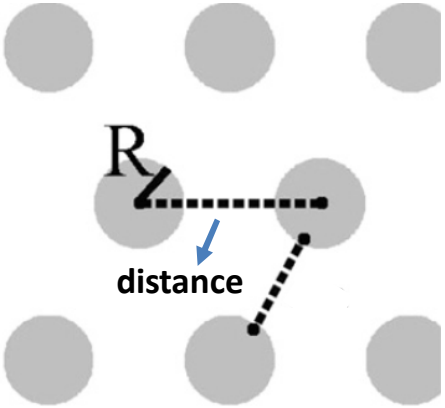


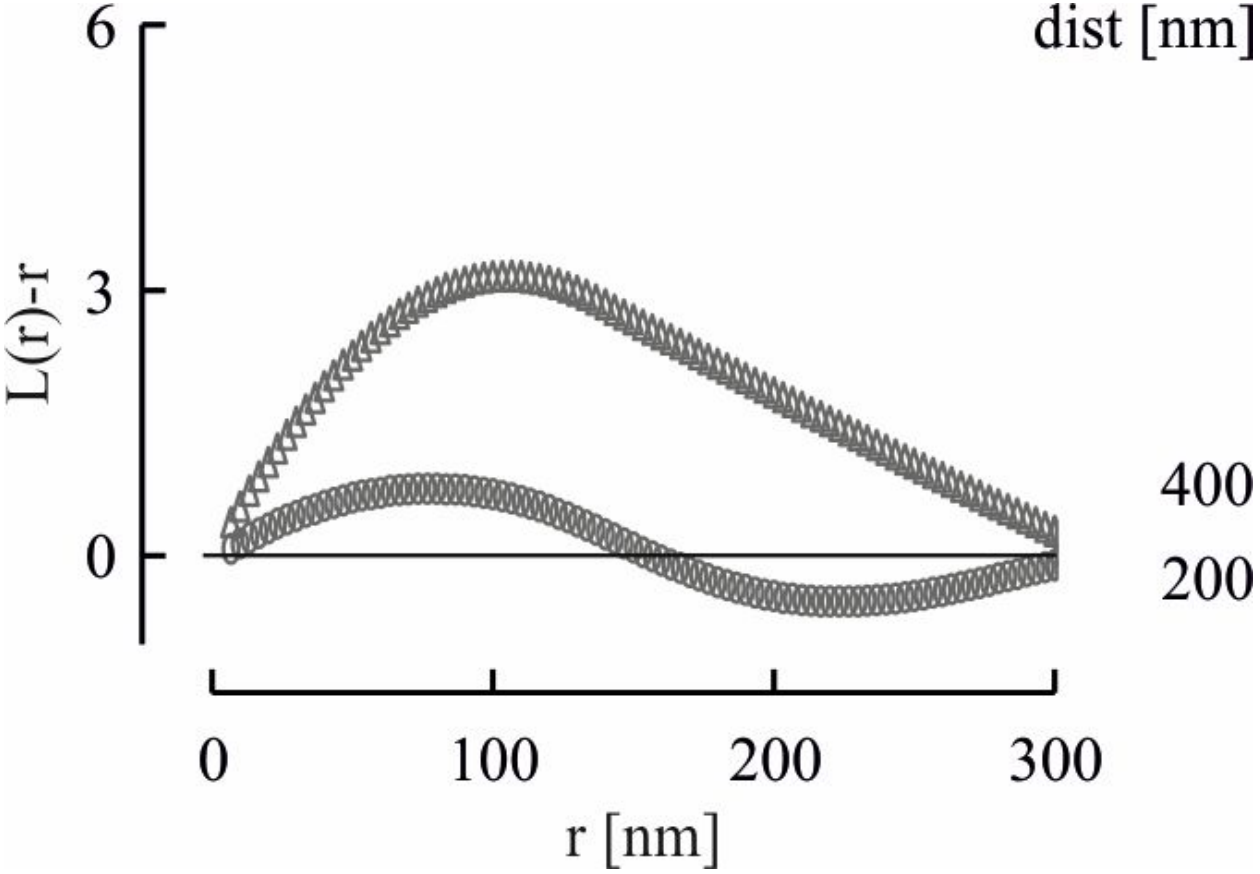


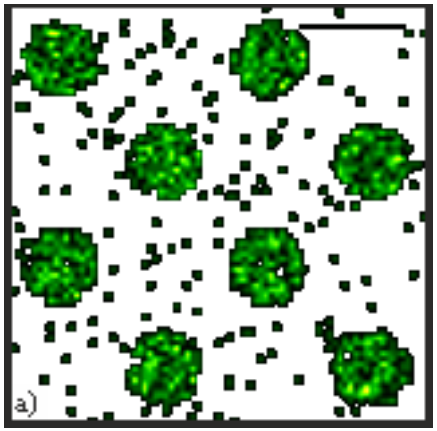
S. Malkusch et al. *Histochem Cell Biol* **139**, 173-9 (2013)

Ripley's functions

Domain size

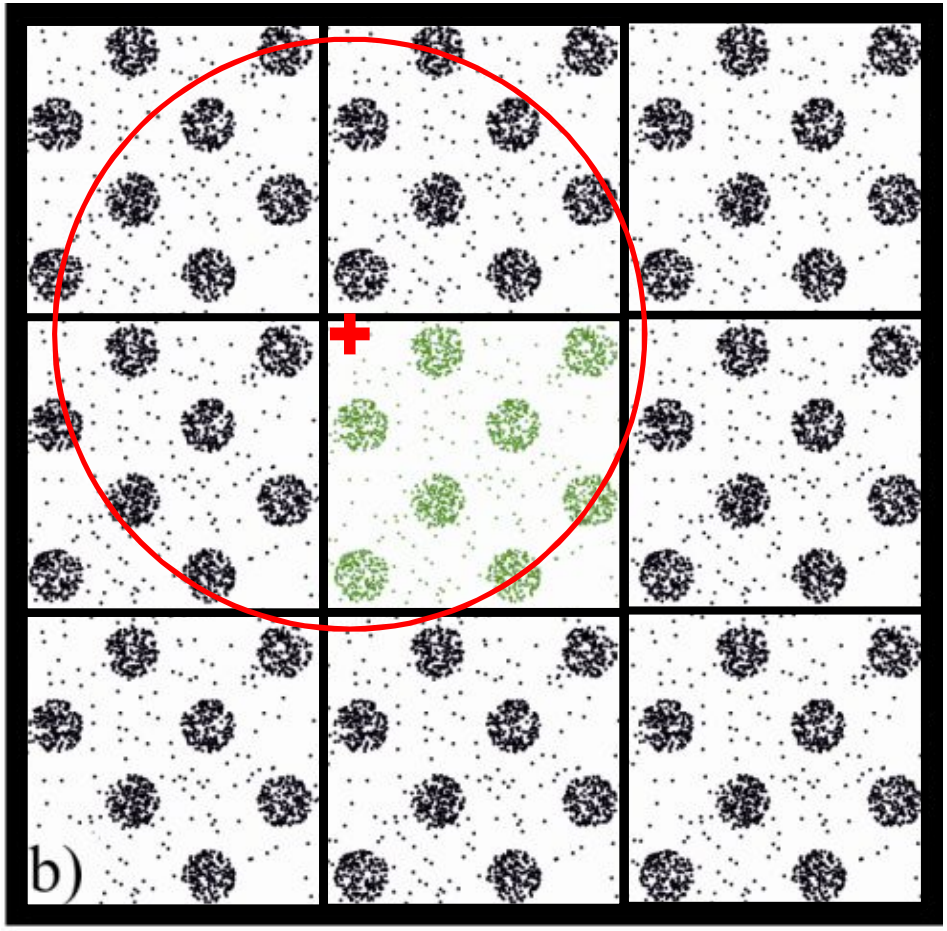






Ripley's functions

Torroidal edge correction

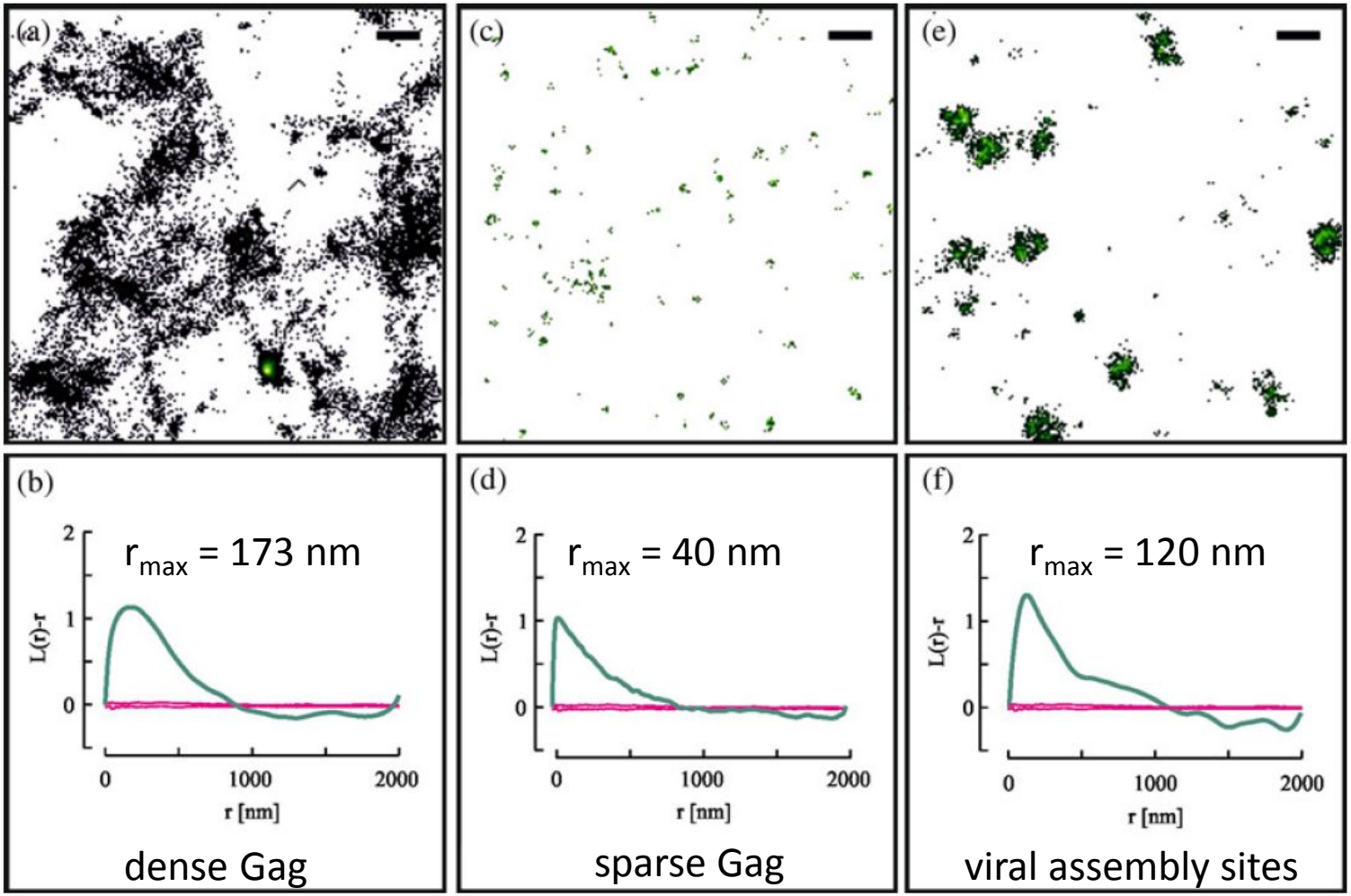


Alternatives: buffer zones, weighted correction

S. Malkusch et al. *Histochem Cell Biol* **139**, 173-9 (2013)

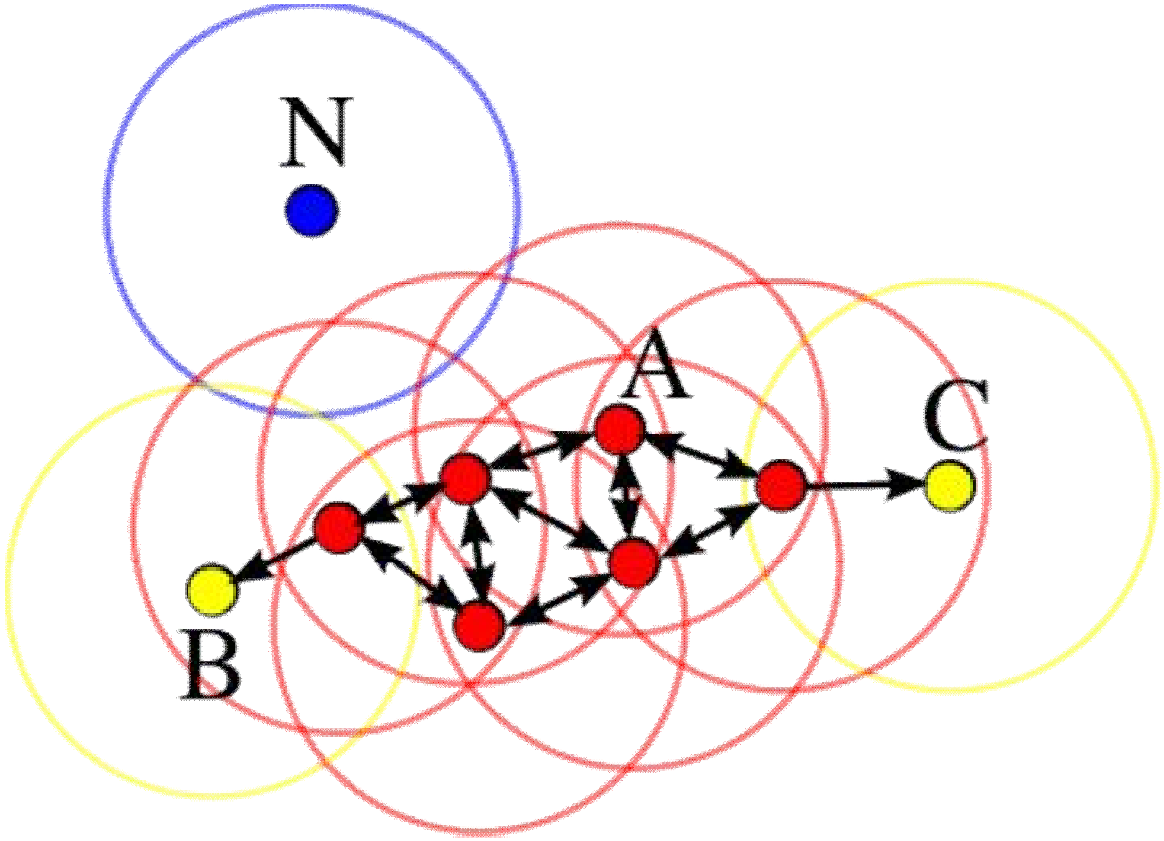
Ripley's functions

Example: Distributions of Gag proteins at the plasma membrane



S. Malkusch et al. *Histochem Cell Biol* **139**, 173-9 (2013)

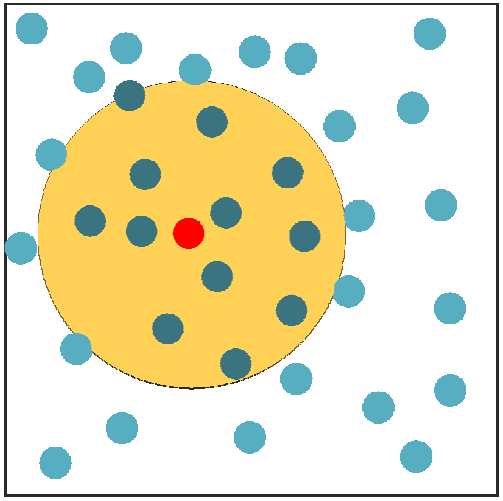
DBSCAN cluster analysis

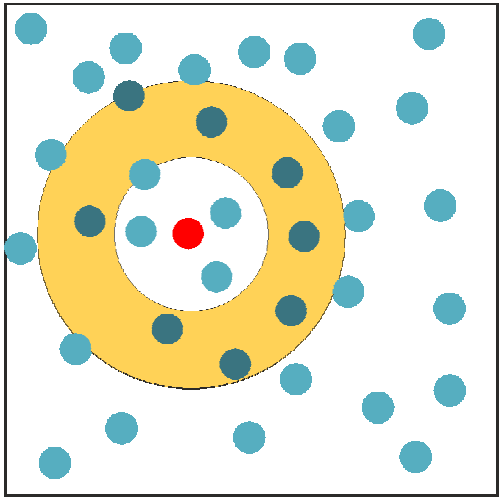


In this diagram, minPts = 3. Point A and the other red points are core points, because at least three points surround it in an ε radius. Because they are all reachable from one another, they form a single cluster. Points B and C are not core points, but are reachable from A (via other core points) and thus belong to the cluster as well. Point N is a noise point that is neither a core point nor density-reachable.

Pair-correlation function

Comparison of Ripley‘s and the pair-correlation function





N = number of localizations in ROI

A = size of ROI

p_i = localization i (here: red point)

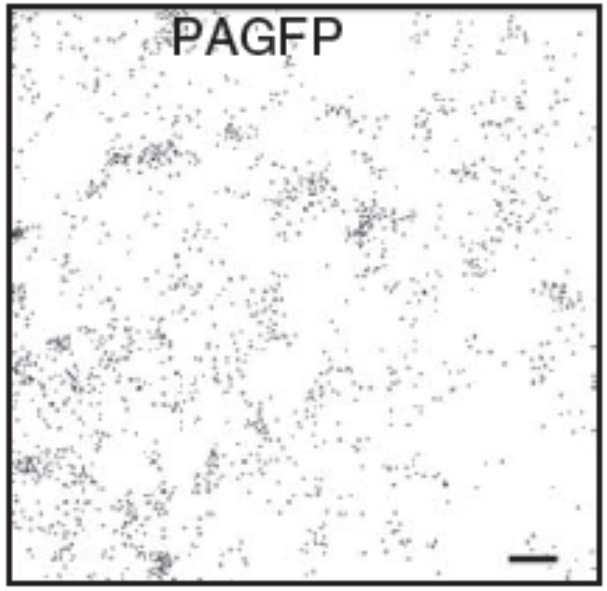
Pair-correlation function

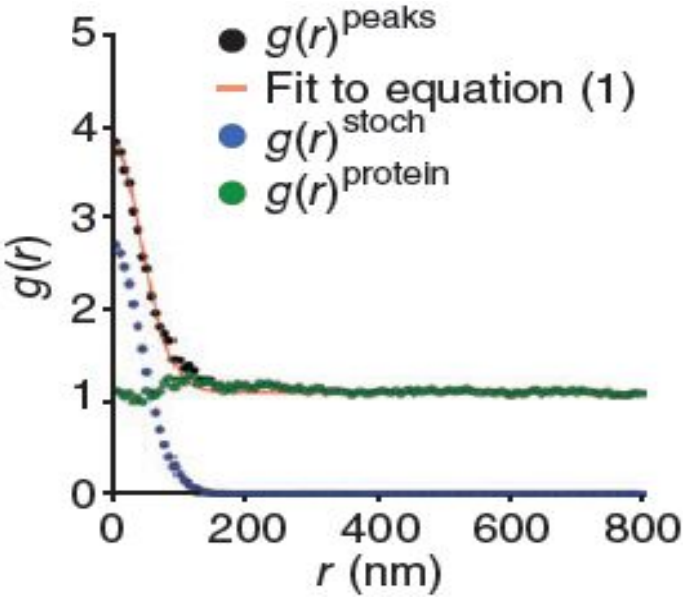
$$g(r) = \frac{A}{4\pi r^2 N^2} \sum_{i=1}^{N} N_{p_i} \; (r_{min} < d \leq r_{max})$$

N_{p_i} = number of locs. around p_i within distance $r_{min} < d \leq r_{max}$

Pair-correlation function

Immobilized fluorophores: complete spatial randomness?





Pair-correlation function fit:

$$g(r) = g_{PSF}(r) \cdot \frac{1}{\rho} + 1 = \frac{1}{4\pi\sigma^2} exp\left(-\frac{r^2}{4\sigma^2}\right) \cdot \frac{1}{\rho} + 1$$

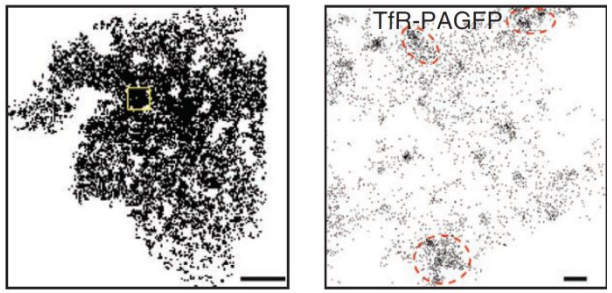
σ = localization precision

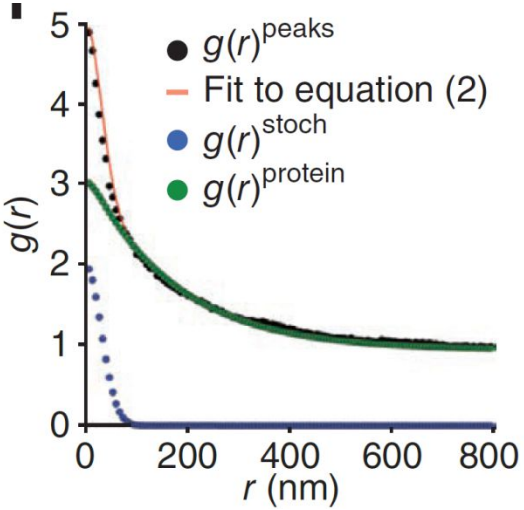
ρ = particle density

P. Sengupta et al. *Nat Meth* **8**, 969-75 (2011)

Pair-correlation function

Example: transferrin receptor labeled with PAGFP





Inhomogeneous protein distribution

Cluster radius $\xi = \mathbf{160\ nm}$

Proteins per cluster $N \approx 2A\pi\xi^2\rho = \mathbf{13}$

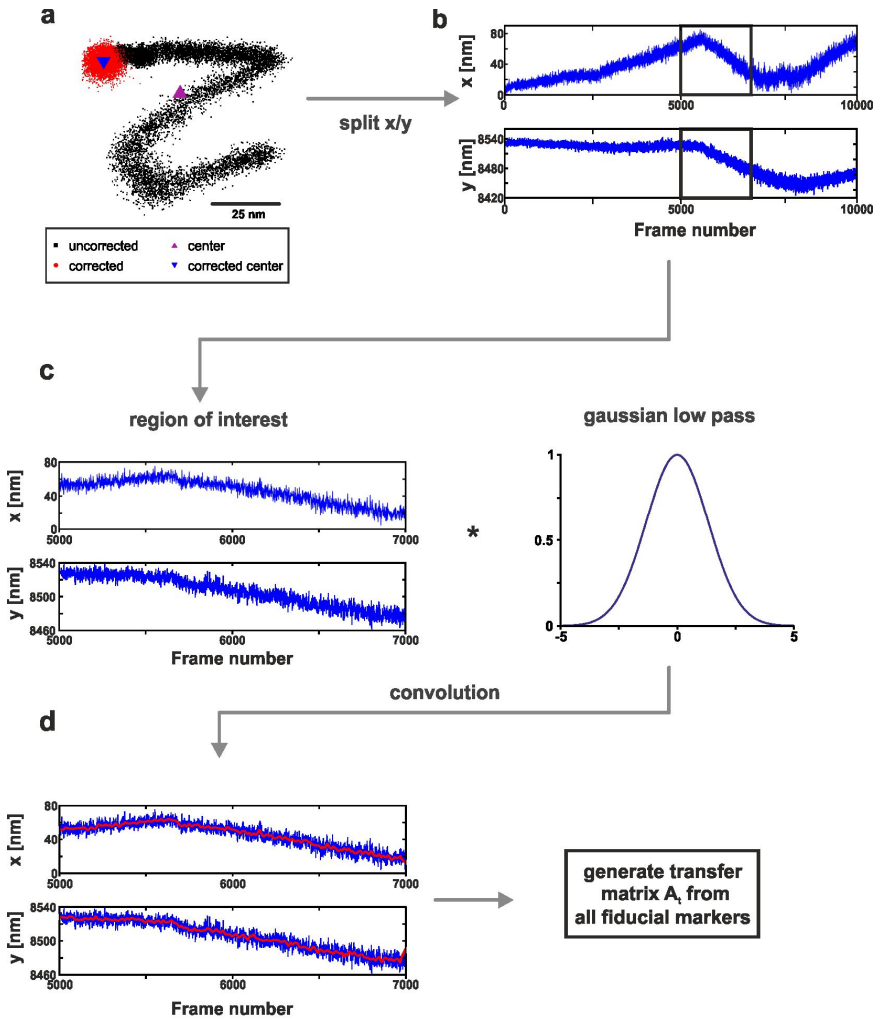
Increase of protein density in clusters $\Psi \approx 2A = \mathbf{3}$

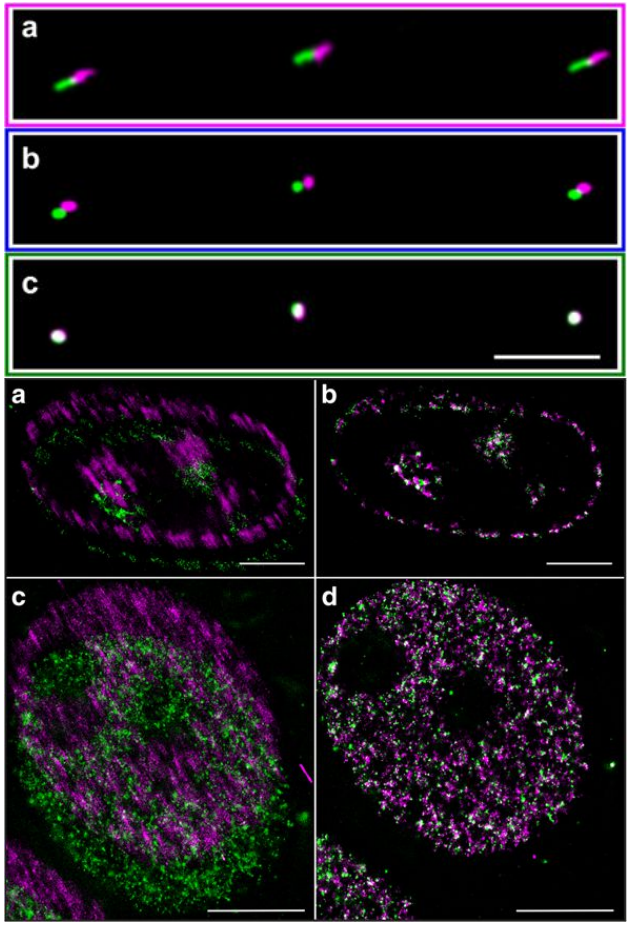
$$g(r) = g_{PSF}(r) \cdot \frac{1}{\rho} + g(r)^{protein} + 1 = \frac{1}{4\pi\rho\sigma^2}\exp\left(-\frac{r^2}{4\sigma^2}\right) + A\exp\left(-\frac{r}{\xi}\right) + 1$$

P. Sengupta et al. *Nat Methods* **8**, 969-75 (2011); P. Sengupta et al. *Chem Rev* **114**, 3189-3202 (2014)

Drift Correction

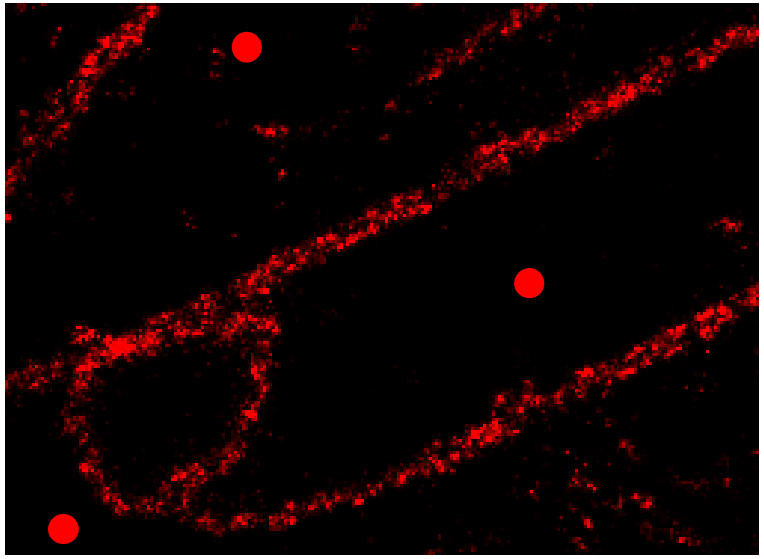
Drift correction with fiducial markers

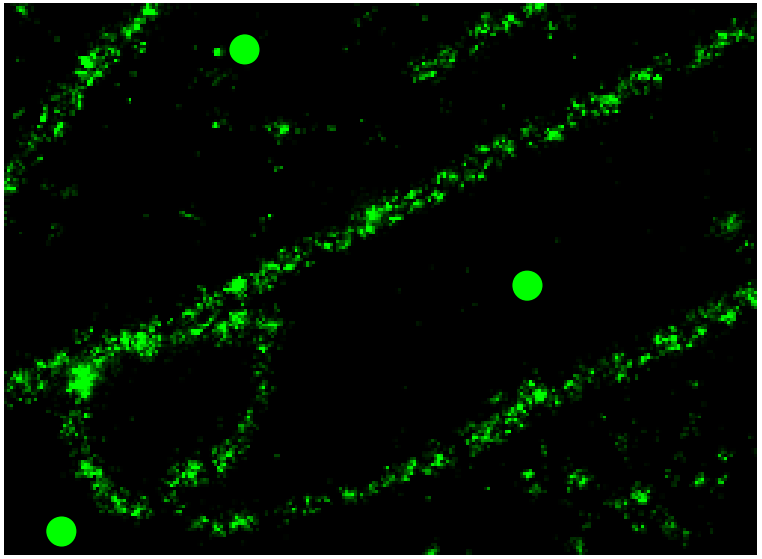




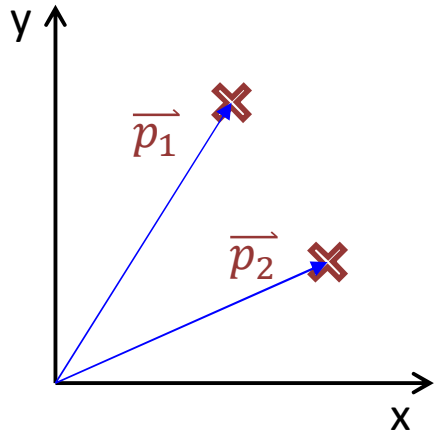
P. Zessin, *Optical Nanoscopy 2*, 1-8 (2013)

Dual-color images





Affine transformation



$$\overrightarrow{p_i} = \begin{pmatrix} p_{x_i} \\ p_{y_i} \end{pmatrix}$$

Translation:

$$\overrightarrow{p_i}' = \overrightarrow{p_i} + \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix}$$

Rotation:

$$\overrightarrow{p_i}' = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix} \overrightarrow{p_i}$$

Scaling:

$$\overrightarrow{p_i}' = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \overrightarrow{p_i}$$

Shearing:

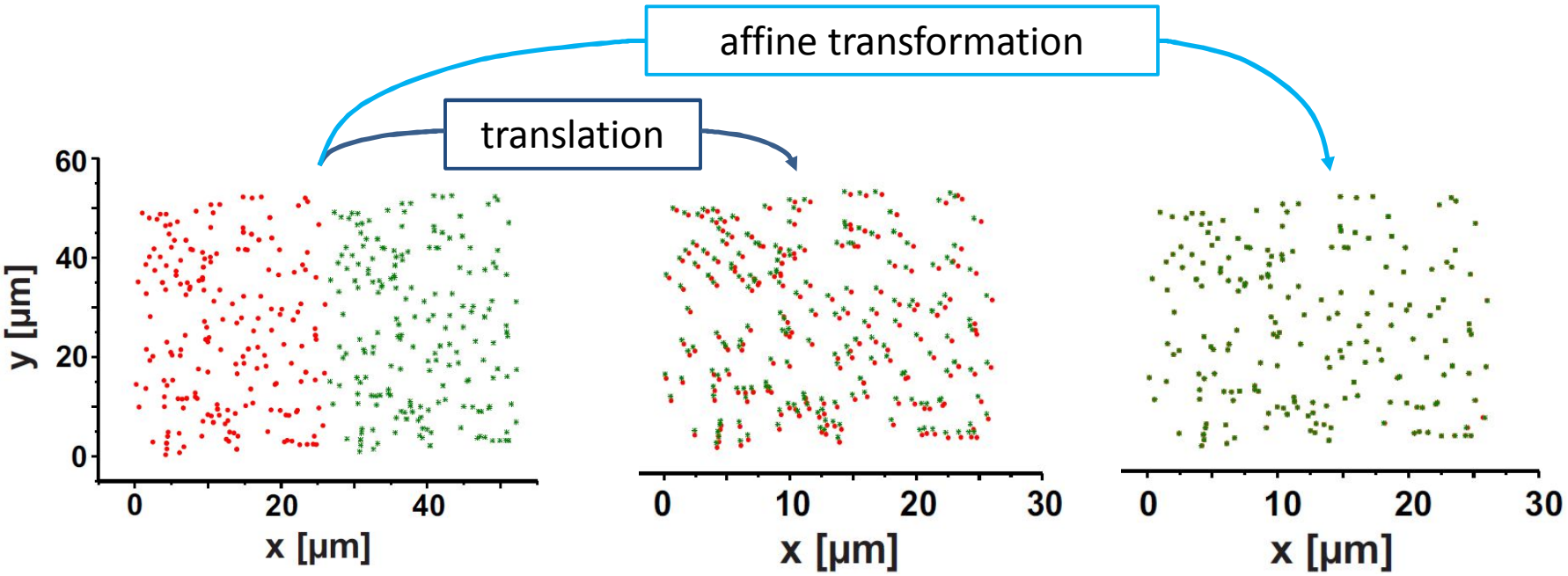
$$\overrightarrow{p_i}' = \begin{pmatrix} 0 & c \\ d & 0 \end{pmatrix} \overrightarrow{p_i}$$

General:

$$\overrightarrow{p_i}' = \boldsymbol{\hat{A}}\, \overrightarrow{p_i} + \vec{\boldsymbol{q}}$$

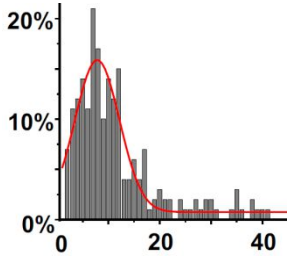
Color-channel alignment

Example: multi-color fluorescent beads in red and green channel



Affine transformation: *translation, rotation, scaling, shearing*

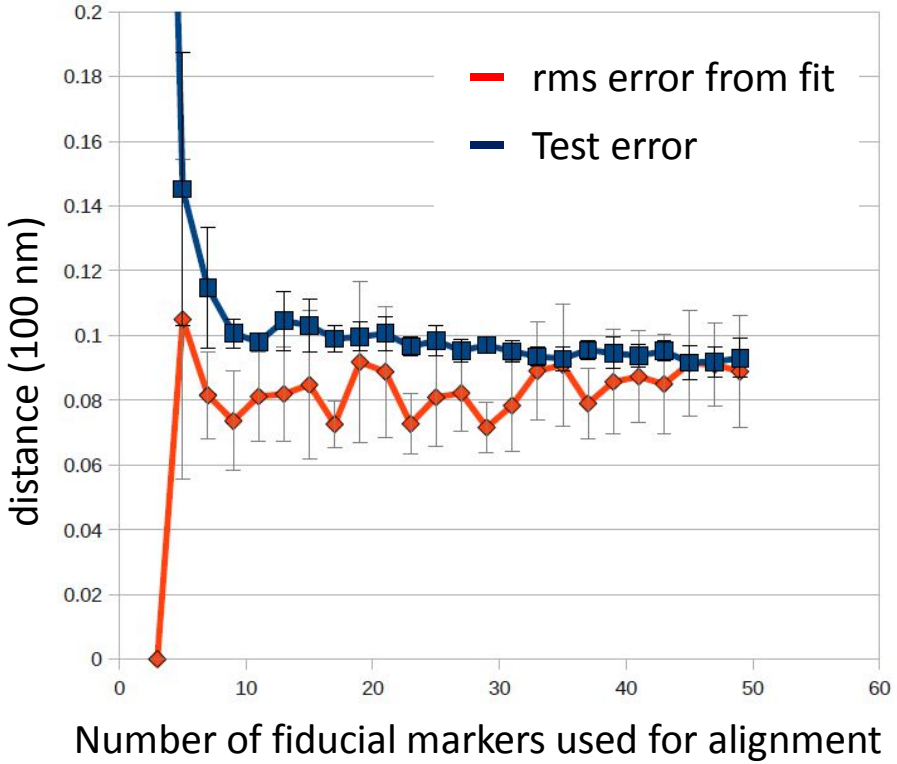
- corrects for chromatic aberration (lenses, mirrors), setup instability (filter change) …
- Here: mean bead displacement error = 7.8 nm



S. Malkusch et al. *Histochem Cell Biol* **137**, 1-10 (2012)

Color-channel alignment

Error of channel-alignment using Tetraspecks



9 fiducial markers are suffiecient to decrease the rms error to ~ 10 nm!

$$\mathrm{rms}(\beta, X, Y) = \sqrt{\frac{1}{N - \mathrm{dof}} \sum_{i=1}^{N} \|x_i - \beta y_i\|^2}$$

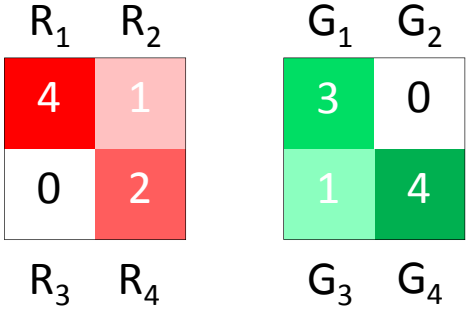
rms: root-mean-square error

J. Schleicher, *Diploma thesis* (2011)

Pearson's correlation coefficient

Describes the degree of overlap between two images

$$r_p = \frac{\sum_i (R_i - R_{aver}) \cdot (G_i - G_{aver})}{\sqrt{\left[\sum_i (R_i - R_{aver})^2 \cdot \sum_i (G_i - G_{aver})^2\right]}}$$



$$R_{aver} = \frac{7}{4} = 1.75$$

$$G_{aver} = \frac{8}{4} = 2$$

$$r_p = \frac{(4-1.75)(3-2) + (1-1.75)(0-2) + (0-1.75)(1-2) + (2-1.75)(4-2)}{\sqrt{[(4-1.75)^2 + (1-1.75)^2 + (0-1.75)^2 + (2-1.75)^2] \cdot [(3-2)^2 + (0-2)^2 + (1-2)^2 + (4-2)^2]}} = 0.64$$

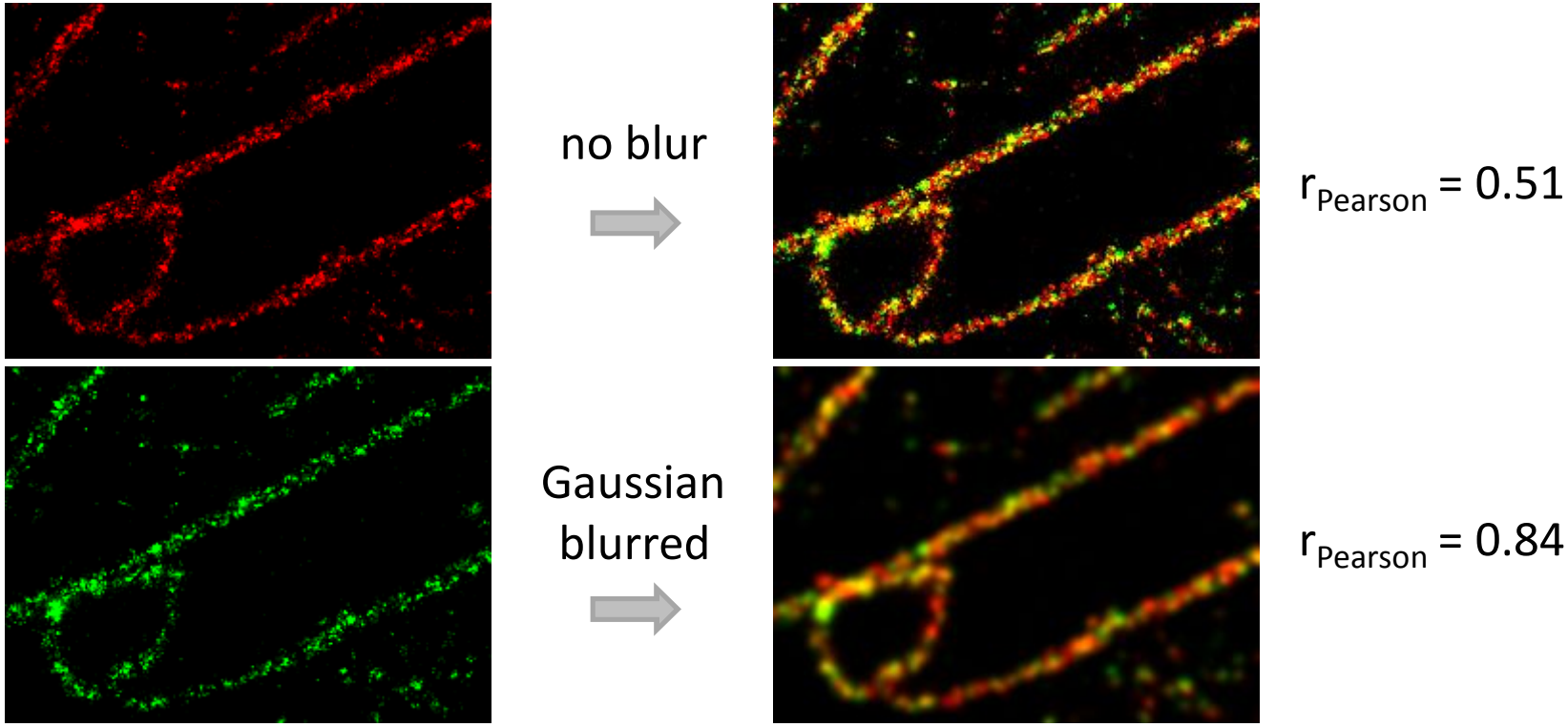
Characteristics of Pearson's coefficient:

- Decides whether patterns correlate in a *linear way*
- not sensitive to the intensity of background or overlapping pixels
- *Range*: 1 ⇢ 0 ⇢ - 1

  perfect correlation ⇢ no correlation ⇢ anti-correlation

E. M. M. Manders et al. *J Microscopy* **169**, 375-82 (1993)

Pearson’s correlation coefficient

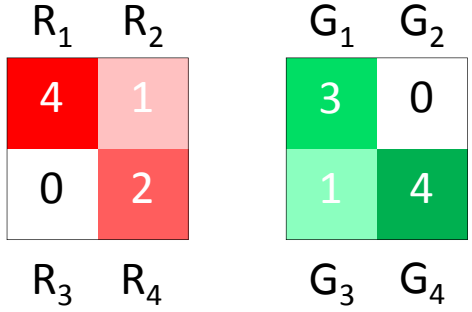


Manders' overlap coefficient

Describes the degree of overlap between two images

$$r = \frac{\sum_i R_i \cdot G_i}{\sqrt{\left[\sum_i (R_i)^2 \cdot \sum_i (G_i)^2\right]}}$$

$$r = \frac{4 \cdot 3 + 1 \cdot 0 + 0 \cdot 1 + 2 \cdot 4}{\sqrt{(4^2 + 1^2 + 0^2 + 2^2)(3^2 + 0^2 + 1^2 + 4^2)}} = 0.64$$



Characteristics of Manders' overlap coefficient :

- proportional to the number of colocalizing objects ($R_i > 0$ and $G_i > 0$)
- Not sensitive to differences in signal intensities
- Ambiguous: number of colocalizing objects have strong influence
- *Range*: 1 ⇢ 0

  perfect correlation ⇢ no correlation

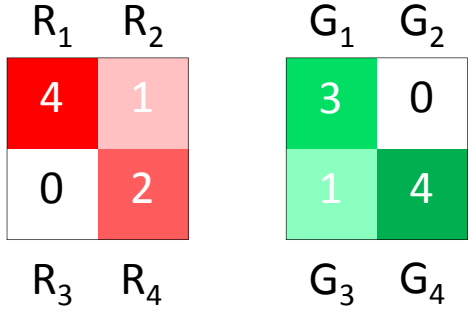
E. M. M. Manders et al. *J Microscopy* **169**, 375-82 (1993)

Manders' coefficients

Describes the degree of overlap between two images

$$M_1 = \frac{\sum_i R_{i,\text{coloc}}}{\sum_i R_i} \quad \Rightarrow \quad M_1 = \frac{4+0+0+2}{4+1+0+2} = 0.86$$

$$M_2 = \frac{\sum_i G_{i,\text{coloc}}}{\sum_i G_i} \quad \Rightarrow \quad M_2 = \frac{3+0+0+4}{4+0+1+4} = 0.88$$



Characteristics of Manders' coefficients :

- equals the number of colocalizing objects/channel (e.g. 86% of red objects colocalize)
- dependent on the signal intensity (background!)
- Not sensitive to differences in signal intensities
- *Range*: 1 ⇢ 0

  perfect correlation ⇢ no correlation

E. M. M. Manders et al. *J Microscopy* **169**, 375-82 (1993)

Costes' method

Setting a background threshold for Manders' coefficients

1. Test colocalization (95% confidence level)

Costes' randomization: test colocalization via Pearson's by comparing with a number of trials, where pixels are "scrambled" in one channel

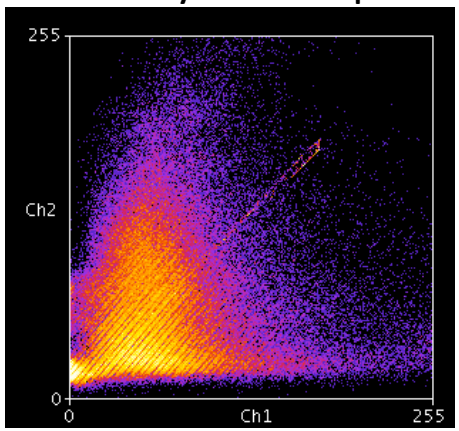
2. Scatter plot

Create a scatter plot of pixel intensities → make linear fit

3. Find threshold for both channels

Find specific point on line, where Pearson's coefficient = 0 for all pixels with values below point → threshold

Intensity scatter plot



S. V. Costes et al. *Biophys J* **86**, 3993-4003 (2004)

Coordinate-based Colocalization (CBC)

Colocalization of single-molecules

Distribution function of neighboring localizations:

$$D_{A_i,\mathrm{A}}(r) = \frac{N_{A_i,\mathrm{A}}(r)}{\pi r^2} \cdot \frac{\pi R_{\max}^2}{N_{A_i,\mathrm{A}}(R_{\max})} = \frac{N_{A_i,\mathrm{A}}(r)}{N_{A_i,\mathrm{A}}(R_{\max})} \cdot \frac{R_{\max}^2}{r^2}$$

$$D_{A_i,\mathrm{B}}(r) = \frac{N_{A_i,\mathrm{B}}(r)}{N_{A_i,\mathrm{B}}(R_{\max})} \cdot \frac{R_{\max}^2}{r^2}$$

Spearman's rank correlation coefficient from distributions

$$S_{A_i} = \frac{\sum_{r_j=0}^{R_{\max}} \left(O_{D_{A_i,\mathrm{A}}}(r_j) - \bar{O}_{D_{A_i,\mathrm{A}}}\right)\left(O_{D_{A_i,\mathrm{B}}}(r_j) - \bar{O}_{D_{A_i,\mathrm{B}}}\right)}{\sqrt{\sum_{r_j=0}^{R_{\max}} \left(O_{D_{A_i,\mathrm{A}}}(r_j) - \bar{O}_{D_{A_i,\mathrm{A}}}\right)^2} \sqrt{\sum_{r_j=0}^{R_{\max}} \left(O_{D_{A_i,\mathrm{B}}}(r_j) - \bar{O}_{D_{A_i,\mathrm{B}}}\right)^2}}$$

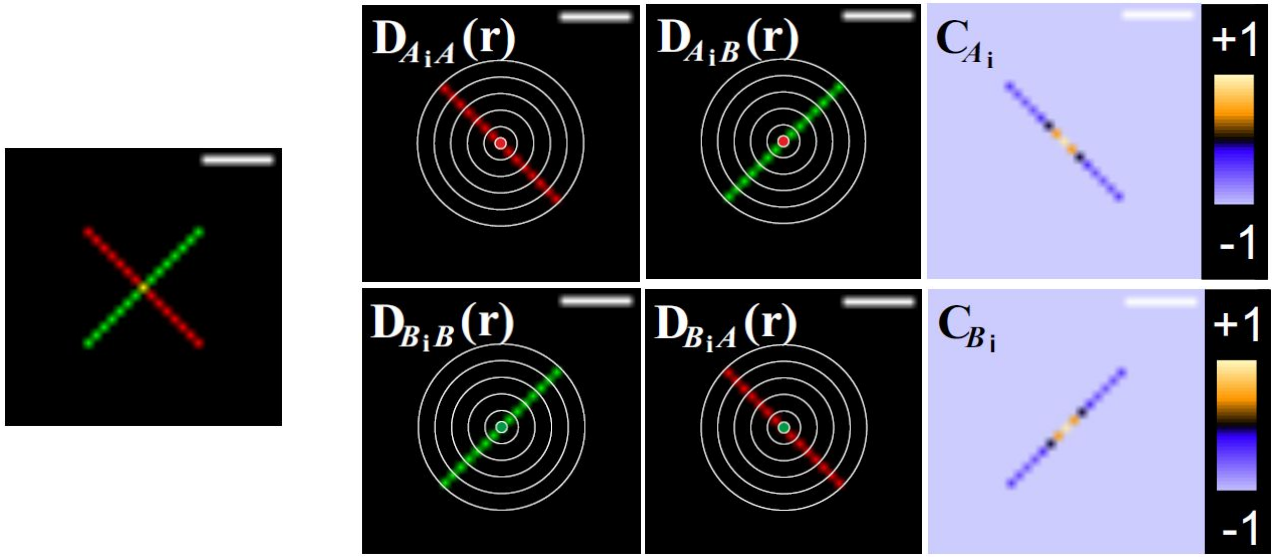
Colocalization value of ith localization of species A/B:

$$CBC_{A_i} = S_{A_i} \exp\left(-\alpha \frac{NN_{A_i,B}}{R_{max}}\right) \qquad CBC_{B_i} = S_{B_i} \exp\left(-\alpha \frac{NN_{B_i,A}}{R_{max}}\right)$$

S. Malkusch et al. *Histochem Cell Biol* **137**, 1-10 (2012)

Coordinate-based Colocalization (CBC)

Colocalization of single-molecules



$$CBC_{A_i} = S_{A_i}\, e^{-\alpha \frac{NN_{A_i,B}}{R_{max}}}$$

Characteristics of CBC:

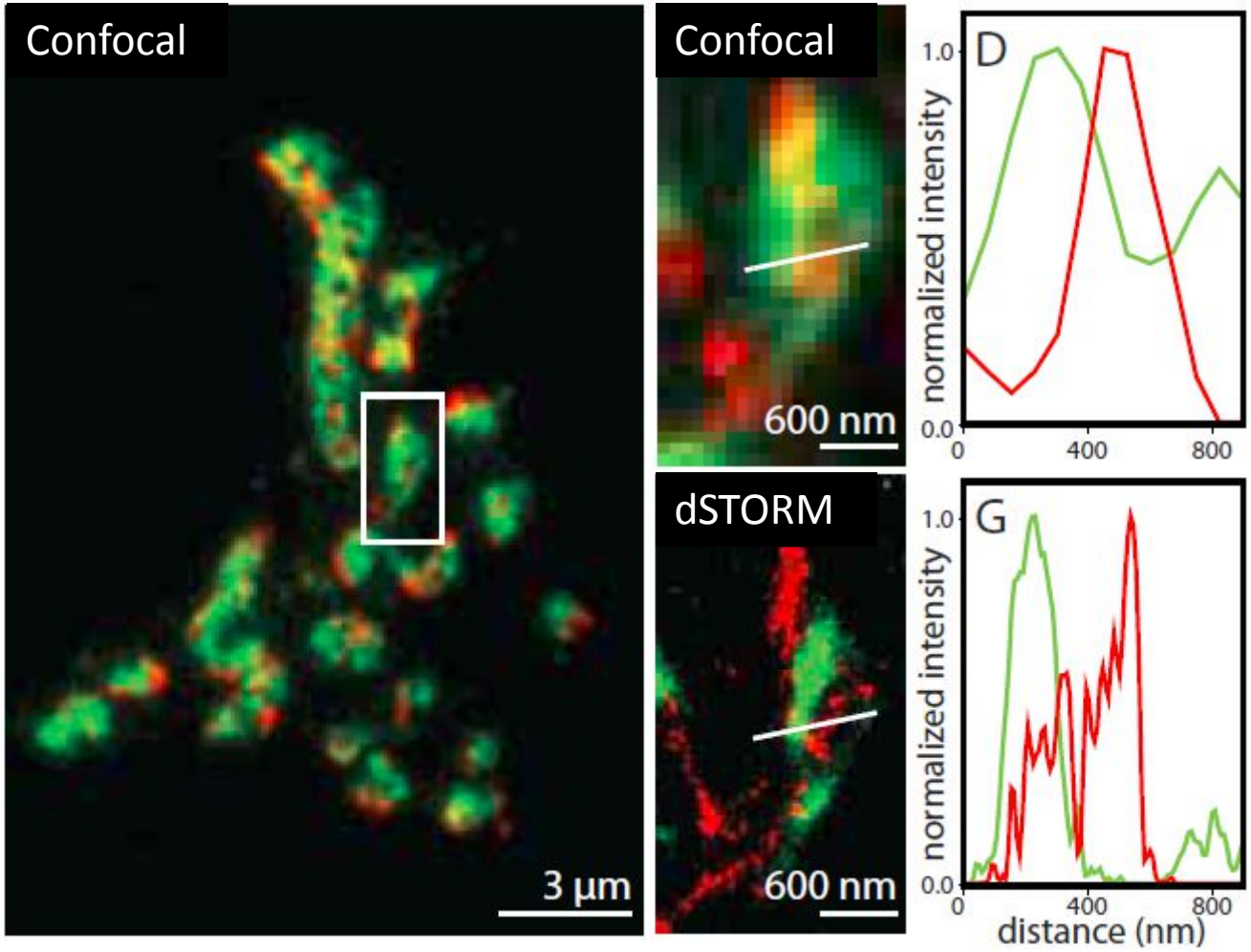
- Colocalization value is assigned to every localization
- Pixel values are averages of colocalization values of single localizations
- *Range*: 1 ⇢ 0 ⇢ - 1

  perfect correlation ⇢ no correlation ⇢ anti-correlation

S. Malkusch et al. *Histochem Cell Biol* **137**, 1-10 (2012)

Image-based Colocalization

Example: cis- (GM130) and trans- (GalT) Golgi membrane proteins



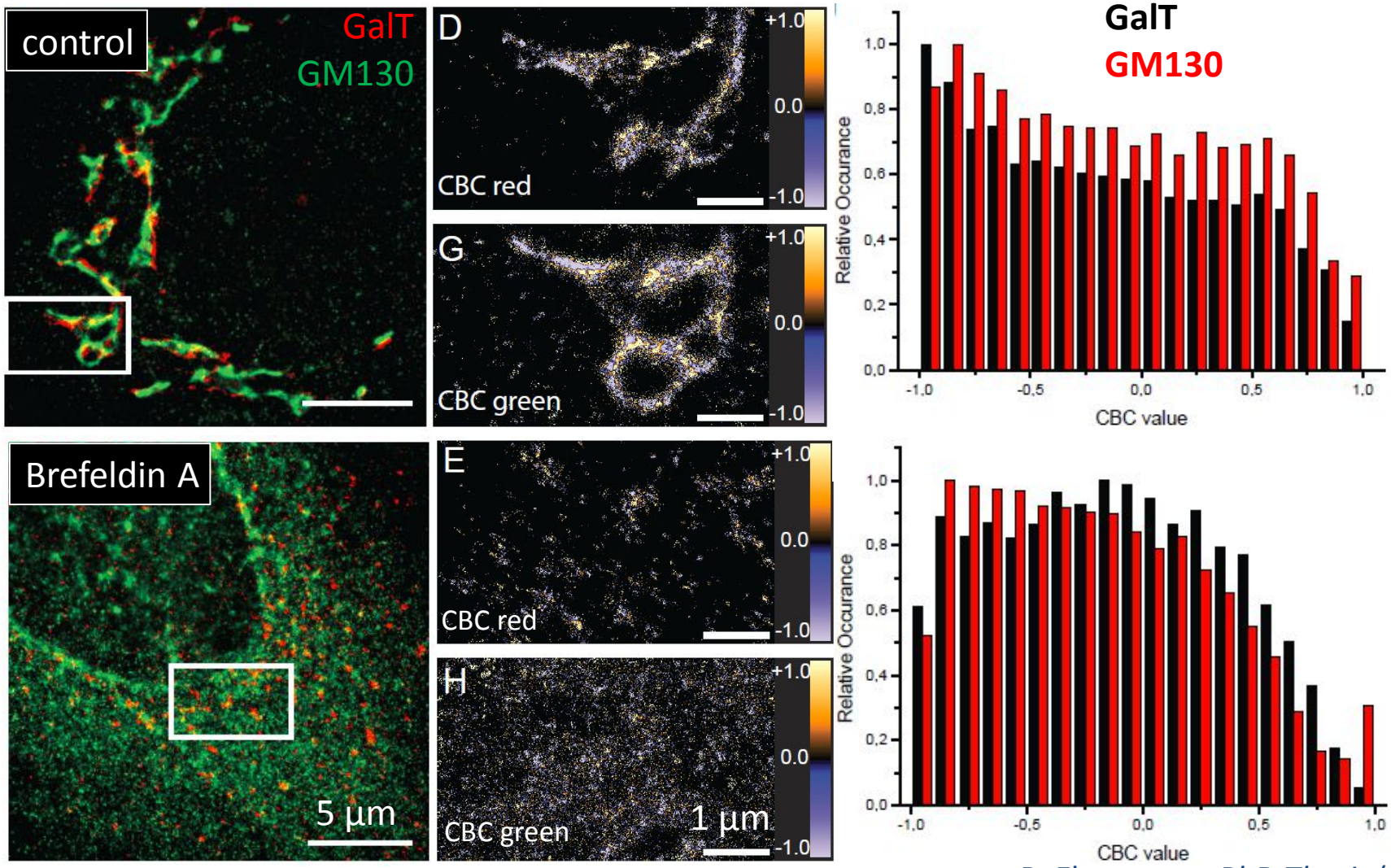
Confocal:
9.6 % colocalization

dSTORM:
5.5 % colocalization

B. Flottmann, *PhD Thesis* (2014)

Coordinate-based Colocalization

Example: cis- (GM130) and trans- (GalT) Golgi membrane proteins



B. Flottmann, *PhD Thesis* (2014)